Journal of Molecular Diagnostics, Vol. 7, No. 1, February 2005

Low Copy Number DNA Template Can Render Polymerase Chain Reaction Error Prone in a Sequence-Dependent Manner

Mansour Akbari,* Marianne Doré Hansen,* Jostein Halgunset,† Frank Skorpen,*† and Hans E. Krokan*

From the Department of Cancer Research and Molecular Medicine and the Department of Laboratory Medicine,† Children's and Women's Health, Norwegian University of Science and Technology, Trondheim, Norway*

Paraffin-embedded tissue is an important source of material for molecular pathology and genetic investigations. We used DNA isolated from microdissected formalin-fixed, paraffin-embedded gastric tumors for mutation analysis of a region of the human gene for uracil-DNA glycosylase (*UNG*), encoding the UNG catalytic domain, and detected apparent base substitutions which, after further investigation, proved to be polymerase chain reaction (PCR) artifacts. We demonstrate that low DNA template input in PCR can generate false mutations, mainly guanine to adenine transitions, in a sequence-dependent manner. One such mutation is identical to a mutation previously reported in the *UNG* gene in human glioma. This phenomenon was not caused by microheterogeneity in the sample material because the same artifact was seen after amplification of a homogenous, diluted plasmid. We did not observe genuine mutations in the *UNG* gene in 16 samples. Our results demonstrate that caution should be taken when interpreting data from PCR-based analysis of somatic mutations using low amounts of template DNA, and that methods used to enrich putative subpopulations of mutant molecules in a sample material could, in essence, be a further amplification of sequence-dependent PCR-generated artifacts. *(J Mol Diagn 2005, 7:36–39)*

Most of the methods for mutation analysis currently available rely on polymerase chain reaction (PCR) amplification of the target DNA. The fidelity of commercially available thermostable DNA polymerases during PCR has been investigated and found to vary considerably.[1,2] The organization sequence and secondary structure of DNA template could also affect the fidelity of the DNA polymerases. Thus, generation of false frameshifts during PCR amplification of microsatellite DNA by different thermostable DNA polymerases has been reported.[3] Such short repetitive DNA sequences are particularly difficult to analyze because of "slippage" of a primer-template complex containing a loop in either the primer or template strand during extension by a DNA polymerase. PCR-based mutation analysis of paraffin-embedded tissues is an important method in cancer diagnostics, and in genetic analysis of tumor cells. However, low amounts and often poor quality of DNA isolated from such material represent a serious obstacle to large-scale mutational studies.[4,5] Human tumors contain a high proportion of C to T transition mutations.[6] Human uracil-DNA glycosylase (UNG) is a major DNA glycosylase that counteracts accumulation of such mutations in the genome.[6] A guanine to adenine mutation in exon III of the gene for UNG resulting in a G143R substitution in human glioma was previously reported.[7] To our knowledge this is the only mutation in the coding region of the *UNG* gene in human cancer cells ever reported, although mutations have been reported in the promoter region.[8] Furthermore, germ line mutations in *UNG* have been reported in the patients suffering from the hyper IgM syndrome.[9] Here we report the apparent detection of the mutation identical to the one described in glioma cells,[7] in gastric tumor cells as well as in normal cells isolated from paraffin-embedded material. However, the mutation turned out to be a PCR artifact. Importantly, we demonstrate that very low copy number of DNA template in the PCR mixture can give rise to this and other false mutations, particularly in GC-rich sequences of the *UNG* gene, and presumably at other genes at similar sequence contexts.

Materials and Methods

Samples

Formalin-fixed, paraffin-embedded tissue specimens from 16 patients with gastric cancer were selected, serially sectioned, and transferred to glass slides. All specimens were evaluated by a pathologist. Normal cells and

Supported by the Research Council of Norway, The Norwegian Cancer Society, The Cancer Fund at St. Olav's Hospital, Trondheim, and the Svanhild and Arne Must Fund for Medical Research.

Accepted for publication August 2, 2004.

Address reprint requests to Hans E. Krokan, Department of Cancer Research and Molecular Medicine, Norwegian University of Science and Technology, N-7489 Trondheim, Norway. E-mail: hans.krokan@medisin.ntnu.no.

Table 1. Oligonucleotide Primers, Size of the PCR Products and the Annealing Temperature Used for Mutation Screening of the Uracil-DNA Glycosylase (*UNG*) gene. All the Reaction Mixtures Were Pre-Treated at 95°C for 10 Minutes (Hot-Start) Prior to Adding DNA Polymerase

Exon/ intron	Primer sequence	Size of amplicons (bp) and GC-content (%)	Annealing temperature and time
IA	5′-ttc gct gcc tcc tca gct cca-´3 5′-ttc ctc ccc ctt cac ccc cta-´3	199 bp 70% GC	63°C 30 seconds
IB	5′-gcc atc cca gcc aag aag-´3 5′-gga agg tgc att tcc ata ttt a-´3	299 bp 63.5% GC	56°C 1 minute
IB	5′-gct ttt gct ggg acc tgt tc-3′ 5′-cca gcc ggg gcc ttc ttg-3′	159 bp 67.9% GC	59°C 30 seconds
II	5′-gag caa tct gat ttt aag tc-´3 5′-tat caa caa gtt gta ctt ac-´3	154 bp 37.4% GC	53°C 30 seconds
III	5′-agg gtc tgt gct gct tac a-´3 5′-aag gaa tca tgt ctg gac tg-´3	200 bp 49.8% GC	55°C 30 seconds
III-s	5′-tta att cct gac ccc tgg tg-´3 5′-gca gtc acc tgt aaa gca ac-´3	151 bp 52.3% GC	55°C 30 seconds
IV	5′-ttc ttg tgg ctt gct ttc ag-´3 5′-aaa atc tag cag tcg ctg gc-´3	133 bp 42.9% GC	57°C 30 seconds
V	5′-ttt aac ctg ttt ctc tca tgt g-´3 5′-gag ccc cag agc aag aaa ac-´3	170 bp 50% GC	55°C 30 seconds
V	5′-agc agt tca ctg atg cag tt-3′ 5′-att ctt tac att gtt ttt ata gtg-3′	167 bp 36.8% GC	55°C 30 seconds
VI	5′-ctg cca cca ttt att ctt ga-´3 5′-agc agc ttc tca aag gcc ac-´3	204 bp 48.8% GC	53°C 30 seconds

tumor cells were microdissected from each slide and transferred to separate Eppendorf tubes. Paraffin was removed by isooctane extraction, and cells were resuspended in 100 μl lysis buffer (50 mmol/L Tris-HCl, pH 8.5, 1 mmol/L EDTA, pH 8.5, 0.5% Tween 20, Acros Organics, Phillipsburg, NJ), 100 μg proteinase K, Sigma Aldrich Chemie, GmbH, Steinham, Germany), and incubated at 42°C overnight to release cellular DNA. Proteinase K was inactivated at 95°C on a PCR block (hot-start PCR) before adding DNA polymerase. Alternatively the DNAs were purified by phenol/chloroform extraction followed by salt/ethanol precipitation. Snap-frozen tissue samples from these patients were not available. Moreover, a P1 phage clone containing the entire human *UNG* gene[10] was used to set up serial dilutions.

PCR and DNA Sequencing

Different parts of the human *UNG* gene were PCR-amplified from cellular DNA and the P1 phage clone, using sets of oligonucleotide primer pairs flanking coding regions of the gene. The sequences of the primers and the corresponding annealing profiles are shown in Table 1. Each PCR was performed for 40 cycles and contained (in a total of 25 μl) one μl cell lysate, 1.5 mmol/L $MgCl_2$, 200 μmol/L of each deoxynucleotide triphosphate (dNTP), 10 mmol/L Tris-HCl (pH 8.3), 5 mmol/L KCl, and one unit *Taq*DNA polymerase (Perkin Elmer, Wellesley, MA). When using *Pfu* DNA polymerase the reaction contained 2 mmol/L $MgSO_4$, 200 μmol/L of each dNTP, 20 mmol/L Tris-HCl (pH 8.8), 10 mmol/L KCl, 10 mmol/L $(NH_4)_2SO_4$, 0.1% Triton X-100, 0.1 mg/ml bovine serum albumin (BSA), and one unit cloned *Pfu* DNA polymerase (Stratagene, La Jolla, CA). The PCR products were purified using a PCR Purification Kit (QIAGEN GmbH, Germany), and sequenced in both directions by cycle sequencing with dye-labeled terminators (BigDye Terminators, PE Applied Biosystems, Foster City, CA), and analyzed on an ABI PRISM 377 DNA Sequencer (PE Applied Biosystems). The DNA sequence of the human *UNG* gene (GenBank Accession No. X89398) was used as reference sequence.

Results

We performed direct sequencing of the PCR-amplified coding region of the *UNG* gene using DNA isolated from paraffin-embedded gastric tumor cells of 16 patients. In nearly 35% of the samples we detected a G to A transition mutation in exon III, identical to a mutation previously reported in a sporadic human glioblastoma[7] (Figure 1). However, repeated analysis of the original samples failed to produce consistent results. Initially, we assumed that a high proportion of wild-type to mutant DNA might override weak signals from the mutant PCR amplicons. Hence, we attempted to enrich for mutant DNA by treating the DNA samples with the endonuclease *Bsm*F1 (New England BioLabs Inc., Beverly, MA) that cleaves the wild-type DNA, but not the mutant DNA, before PCR. We failed to synthesize PCR products from these samples even after 45 cycles of PCR (data not shown). This indicates that the samples predominantly consisted of the wild-type *UNG* sequence, and that the putative mutation most probably was a PCR artifact. We next examined a possible impact of the concentration of DNA template used in the PCR mixture on the generation of this mutation. We found that using three- to fourfold more DNA in the PCR mixture, relative to the initial amount, the final products only consisted of the wild-type DNA molecules as assessed by DNA sequencing (data not shown). Next, we performed PCR on a set of serial dilutions of a

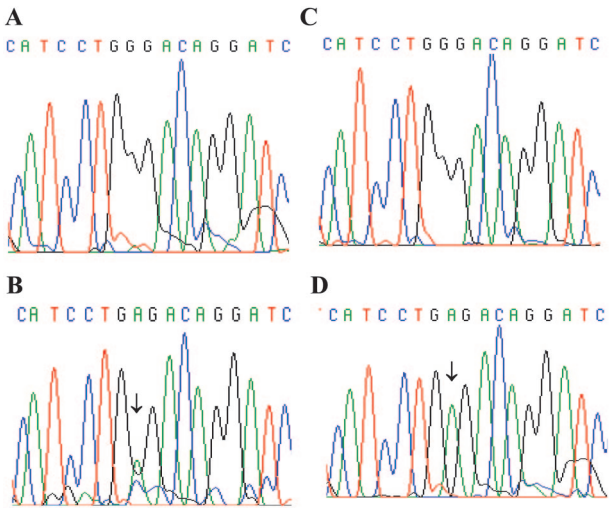

Figure 1. Sequencing analysis of PCR amplified genomic DNA of a patient with gastric cancer. DNA was isolated from morphologically normal and cancer cells of paraffin-embedded tissues. The nucleotides surrounding the reported G to A transition mutation in exon III of the *UNG* gene are shown. **A:** Wild-type sequence. **B:** Position of the G to A transition mutation shown by an **arrow**. We prepared serial dilutions of P1 phage and performed 10 independent PCRs for each titer. The approximate number of P1 vector used in each PCR mix was 50 copies (**C**), 25 copies (**D**). None of the 10 PCR products from the 50-copy titer showed the base substitution. **D:** Represents 6 of 10 PCR reactions (60%).

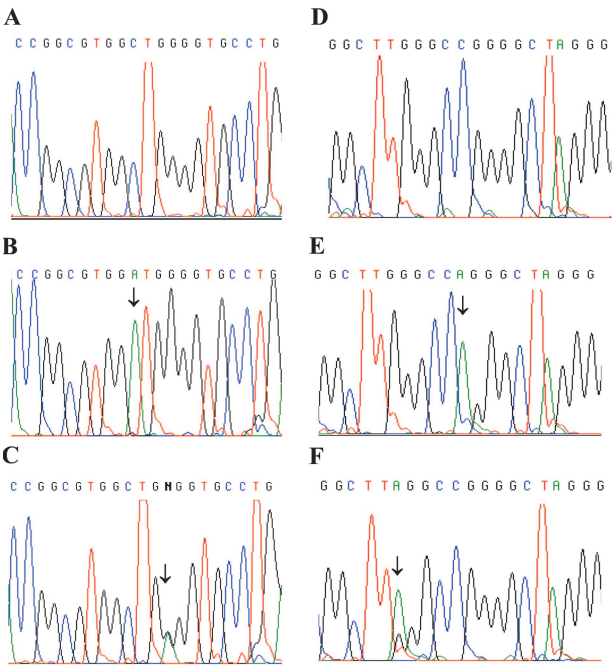

Figure 2. A–C: A region of the GC-rich exon IA that displayed high incidence of PCR-produced base substitutions at low DNA input. **A:** Wild-type sequence. In **B** and **C**, approximately 25 copies of the P1 phage were used as template in the PCR mixture. **D–F:** A GC-rich region in intron IA that showed high rate of PCR-induced G to A base substitution at very low DNA template concentration. **D:** Wild-type DNA sequence. In **E** and **F,** each PCR mix contained approximately 25 copies of P1 phage DNA template. The **arrows** show the site of base changes.

P1 phage clone, which contains the entire *UNG* gene. We routinely performed 10 independent PCR for each titer. We directly sequenced the PCR products in both directions and found that the apparent occurrence of the G to A transition mutation in exon III was inversely related to the amount of DNA input in PCR mixture (Figure 1). Using a different set of oligonucleotide primers to produce shorter amplicons (Table 1, Exon III-s), as well as using the high-fidelity *Pfu* DNA polymerase in PCR instead of *Taq*DNA polymerase did not eliminate the production of false mutations (data not shown). Next, we searched for PCR-produced base substitutions in other regions of the *UNG* gene, using titers of the P1 phage clone as template. In addition to the initial 10 independent PCR for each sample, we performed five additional PCR on samples that showed high incidence of site-specific PCR errors. At 25 copies of DNA template input in 15 independent PCR reactions of GC-rich exon/introns IA and IB, 40% and 25% contained site-specific PCR artifacts, respectively. The GC content of the amplicons is shown in Table 1. Also here the majority of the base substitutions were G to A transitions. Figure 2 shows selected positions that most frequently displayed occurrence of PCR-induced mutations.

Most of the paraffin-embedded specimens contained a limited number of cancer cells surrounded by normal cells. To avoid contamination of cancer cells with normal cells, we scraped only cancer cells from the middle of tumors, usually numbering less than 50 cells. Given the poor quality of DNA isolated from paraffin-embedded specimens, the amount of amplifiable DNA template roughly estimated corresponds to 25 copies of P1 phage. We did not detect reproducible and site-specific increase of PCR errors in the remaining regions (data not shown).

Discussion

PCR is central in most mutation analysis methods. Commercially available thermostable DNA polymerases routinely used in PCR display different fidelities.[1,2] Although to a lesser extent, the impact of other PCR components such as salt, dNTPs, oligonucleotide primers, as well as the effect of regionally stable secondary structures of DNA template on the fidelity of thermophilic DNA polymerases have also been studied. In the present study, we show that very low concentration of DNA template in PCR mix can give rise to false base substitutions in a sequence-dependent fashion. Importantly, the false mutations detected in the present study was observed in unique sequences, not in repetitive sequences as found in microsatellites. Using the high fidelity *Pfu* DNA polymerase, which possesses 3′→5′ exonuclease (proofreading) activity, did not eliminate the problem. In a previous study, the proofreading activity of *Pfu* DNA polymerase was found not to reduce the production of false frameshifts during amplification of microsatellites.[3] Apparently, the formation of stable secondary structures in DNA could negatively affect the accuracy of DNA polymerases and disable their proofreading function.[3,11] However, this cannot explain the low DNA template concentration-dependent production of such base substitutions. PCR amplification of DNA isolated from formalin-fixed and paraffin-embedded specimens has been reported to be error prone and thought to be a result of base damages and large-scale DNA fragmentation

caused by the chemical preparation of such specimens.[12,13] An early report demonstrated that, when *Taq*DNA polymerase encountered a damaged template, it could "jump" to another template and continue from there.[14] Moreover, *Taq*DNA polymerase was suggested to insert an A instead of a G opposite damaged cytosine nucleotides the so-called "A-rule", hence resulting in artificial G to A mutations.[12] However, our results demonstrate that low copy number DNA template gives rise to PCR errors in a sequence-specific fashion in intact plasmid DNA by *Taq* as well as *Pfu* DNA polymerases. This suggests that DNA damage caused by formalin-fixation is not the sole explanation for the generation of such errors.

The results presented in this work might have important implications for the general application of laser-capture microdissection in mutation studies, a method used to isolate single cells from paraffin-embedded tissue slices.[15] Furthermore, the sequence dependency of PCR errors seen in this study shows that caution must be exercised in the interpretation of clustering of mutations in the regions assumed to represent mutational hotspots.[16,17] PCR-based enrichment of target DNA must be carefully designed and, if possible, supplementary methods such as restriction digestion of template DNA should be included in mutation studies.

References

1. Andre P, Kim A, Khrapko K, Thilly WG: Fidelity and mutational spectrum of Pfu DNA polymerase on a human mitochondrial DNA sequence. Genome Res 1997, 7:843–852
2. Cline J, Braman JC, Hogrefe HH: PCR fidelity of Pfu DNA polymerase and other thermostable DNA polymerases. Nucleic Acids Res 1996, 24:3546–3551
3. Hite JM, Eckert KA, Cheng KC: Factors affecting fidelity of DNA synthesis during PCR amplification of d(C-A)n.d(G-T)n microsatellite repeats. Nucleic Acids Res 1996, 24:2429–2434
4. Cui X, Feiner H, Li H: Multiplex loss of heterozygosity analysis by using single or very few cells. J Mol Diagn 2002, 4:172–177
5. Farrand K, Jovanovic L, Delahunt B, McIver B, Hay ID, Eberhardt NL, Grebe SK: Loss of heterozygosity studies revisited: prior quantification of the amplifiable DNA content of archival samples improves efficiency and reliability. J Mol Diagn 2002, 4:150–158
6. Krokan HE, Drablos F, Slupphaug G: Uracil in DNA: occurrence, consequences, and repair. Oncogene 2002, 21:8935–8948
7. Moon YW, Park WS, Vortmeyer AO, Weil RJ, Lee YS, Winters TA, Zhuang Z, Fuller BG: Mutation of the uracil DNA glycosylase gene detected in glioblastoma. Mutat Res 1998, 421:191–196
8. Kvaloy K, Nilsen H, Steinsbekk KS, Nedal A, Monterotti B, Akbari M, Krokan HE: Sequence variation in the human uracil-DNA glycosylase (UNG) gene. Mutat Res 2001, 461:325–338
9. Imai K, Slupphaug G, Lee WI, Revy P, Nonoyama S, Catalan N, Yel L, Forveille M, Kavli B, Krokan HE, Ochs HD, Fischer A, Durandy A: Human uracil-DNA glycosylase deficiency associated with profoundly impaired immunoglobulin class-switch recombination. Nat Immunol 2003, 4:1023–1028
10. Haug T, Skorpen F, Kvaløy K, Eftedal I, Lund H, Krokan HE: Human uracil-DNA glycosylase gene: sequence organization, methylation pattern, and mapping to chromosome 12q23–q24.1. Genomics 1996, 36:408–416
11. Loewen PC, Switala J: Template secondary structure can increase the error frequency of the DNA polymerase from Thermus aquaticus. Gene 1995, 164:59–63
12. Williams C, Ponten F, Moberg C, Soderkvist P, Uhlen M, Ponten J, Sitbon G, Lundeberg J: A high frequency of sequence alterations is due to formalin fixation of archival specimens. Am J Pathol 1999, 155:1467–1471
13. Quach N, Goodman MF, Shibata D: In vitro mutation artifacts after formalin fixation and error prone translesion synthesis during PCR. BMC Clin Pathol 2004, 4:1 http://www.biomedcentral.com/1472-6890/4/1
14. Paabo S, Irwin DM, Wilson AC: DNA damage promotes jumping between templates during enzymatic amplification. J Biol Chem 1990, 265:4718–4721
15. Dillon D, Zheng K, Costa J: Rapid, efficient genotyping of clinical tumor samples by laser-capture microdissection/PCR/SSCP. Exp Mol Pathol 2001, 70:195–200
16. Fenoglio-Preiser CM, Wang J, Stemmermann GN, Noffsinger A: TP53 and gastric carcinoma: a review. Hum Mutat 2003, 21:258–270
17. Wright BE, Reimers JM, Schmidt KH, Reschke DK: Hypermutable bases in the p53 cancer gene are at vulnerable positions in DNA secondary structures. Cancer Res 2002, 62:5641–5644